

Appendix

Appendix A1.1 Study characteristics: Saunders & Goldenberg, 1999 (randomized controlled trial with reservations)

Characteristic	Description
Study citation	Saunders, W. M., & Goldenberg, C. (1999). Effects of instructional conversations and literature logs on limited- and fluent-English-proficient students' story comprehension and thematic understanding. <i>Elementary School Journal</i> , 99(4), 279–301.
Participants	All 138 English language learners enrolled in three fifth-grade and two fourth-grade classrooms in one school participated in this study. They were matched by language proficiency (limited or fluent) and teachers' rating of reading skills, and then randomly assigned to four study conditions within each classroom. Twenty-two English language learners were excluded from the final analysis (3 special education students, 4 students enrolled just prior to the study, 12 students not present for some of the activities, and 3 students randomly excluded to provide for a balanced design). The remaining 116 participants were evenly distributed among the four groups (29 per group). Only 32 of these participants in two study groups are of interest for this intervention report: 16 English language learners in the comparison condition and 16 English language learners in the <i>Instructional Conversations</i> and <i>Literature Logs</i> (intervention) condition. ¹ These students participated in the study with students fluent in English.
Setting	The study took place in a K–5 elementary school located in an urban area. The majority of students at the school were Hispanic (82%), had limited English proficiency (69%), and qualified for the free or reduced-price lunch program (62%). More than three quarters of the fourth-grade students were performing below grade level in reading, language, and math. The school ranked among the lowest 20% of schools in the district. Schoolwide efforts were underway at the time of the study to improve bilingual programs, English language development program, language arts instruction, and overall academic infrastructure at the school.
Intervention	The study lasted for a period of 10–15 days, including pre- and post-intervention activities. The intervention was implemented over four days. On the first day, English language learners in the <i>Literature Log</i> only group and the <i>Instructional Conversations and Literature Logs</i> group received instruction on <i>Literature Logs</i> in two consecutive 45-minute lessons. On the second day, English language learners in the <i>Instructional Conversation</i> only group and the <i>Instructional Conversations and Literature Logs</i> group received two 45-minute lessons on <i>Instructional Conversations</i> . The same procedures were followed on the third and fourth days, with the order of the lessons reversed (that is, <i>Instructional Conversations</i> lessons on day three and <i>Literature Logs</i> lessons on day four). English language learners in each group also spent at least 45 minutes on creating an illustration and caption summarizing their interpretation of the story “Louella’s Song.”
Comparison	English language learners in the comparison group participated in reading and writing activities related to social studies either independently or with a teaching assistant. They also devoted at least 45 minutes to creating an illustration and caption summarizing their interpretation of the story “Louella’s Song.” Saunders and Goldenberg (1999) reported that English language learners in the comparison group did not receive as much direct instructional time from teachers as those in the <i>Instructional Conversations and Literature Logs</i> group. ²
Primary outcomes and measurement	The effects of the intervention were assessed using several measures, including factual comprehension, interpretive comprehension, theme-explanation essay, and theme-exemplification essay. Although theme-explanation and theme-exemplification essays were measures used in the study, outcomes were not reported in the WWC report because the data reported in the study were the percentage of English language learners whose essays received a high score. Therefore, the measures did not meet WWC standards (see Appendix A2.1 for more detailed descriptions of outcome measures).
Teacher training	The five teacher participants in the study were members of the research and development team implementing the school’s language arts model in Spanish, transition, and mainstream English language arts classrooms. The team was led by two instructional advisors who were able to co-teach and provide assistance in the classrooms on a daily basis. The team met on a bi-monthly basis to study instructional components, view videotapes and live demonstrations, plan instructional units, and evaluate English language learners’ work. All study conditions (including the control) were carried out in each classroom by each of the five teachers.

1. The other two study groups (*Instructional Conversations* only and *Literature Logs* only) in Saunders and Goldenberg (1999) were not of interest to this intervention report.

2. The different amounts of teacher-led instruction in the intervention and comparison conditions might have confounded the effects of the intervention; therefore this study was downgraded to meeting WWC evidence standards with reservations.

Appendix A1.2 Study characteristics: Saunders, 1999 (quasi-experimental design)

Characteristic	Description
Study citation	Saunders, W. M. (1999). Improving literacy achievement for English learners in transitional bilingual programs. <i>Educational Research and Evaluation</i> , 5(4), 345–381.
Participants	Eligible participants for the study were students who enrolled in their respective schools from first grade through fifth grade, considered Spanish their primary home language, had limited English proficiency at the time of enrollment, and were in a Spanish transitional bilingual program. Among the pool of eligible English language learners, 180 were randomly selected at the beginning of second grade from five intervention schools and five neighboring schools, then matched on academic achievement, percentage of students with limited English proficiency, percentage of students in the free or reduced-priced lunch program, school enrollment, and ethnic composition. Of the 180 English language learners, 125 remained at their respective school through fifth grade, from which 42 intervention-group English language learners and 42 comparison-group English language learners matched on first grade Spanish reading and language and transition status were selected for data analysis.
Setting	The study took place in a school district in southern California. The average enrollment for the 10 participating schools was 880. Over 90% of the students enrolled in these schools were Latino or Hispanic, 74% were limited English proficient, and 95% qualified for free or reduced-price lunches.
Intervention	English language learners in the intervention group participated in a transitional bilingual program, which covered several transitional phases as they progressed from receiving instruction in their primary language to instruction in English. English language learners were in the Pre-Transition phase in the second and third grades, in Transition I in the fourth grade, and in Transition II in the fifth grade. Transition II is the only phase during which instruction was dedicated completely to reading and writing in English, and English language learners were assessed with English language measures. Transition II services were provided only in grade 4 or 5, depending on whether a student was ready for English-only instruction. The earlier transition phases involved academic support and assessments in both English and Spanish. This transition program incorporated 12 instructional components, including <i>Instructional Conversations</i> and <i>Literature Logs</i> . <i>Instructional Conversations</i> involved ongoing discussions, while <i>Literature Logs</i> (and writing projects) entailed writing activities. Teachers assisted English language learners in understanding how the experiences of a character in a story relate to their own experiences. Writing assignments were intended to encourage English language learners to think about and express ideas, as well as extract meaning from the stories. The <i>Instructional Conversations</i> were intended to provide a forum for sharing ideas and a more complex understanding of story themes. English language learners studied literature during each phase, practiced using reading comprehension strategies, and read assigned books independently. The assignments were based on each English language learner's reading level. Additionally, for 45 minutes per day during the Pre-Transition phase, teachers provided instruction using the English Language Development through Literature program. Instruction was provided in small groups comprised of English language learners with similar levels of English language proficiency. Three or more times per week, teachers read aloud for approximately 20 minutes. English language learners also chose a book or story to read independently each day and completed short assignments for the ones that they found most engaging. Additionally, they wrote in journals during the beginning stages of English writing and received written responses from the teacher. <i>Instructional Conversations</i> and <i>Literature Logs</i> were two of the 12 components of the language arts program that was implemented.
Comparison	English language learners in the comparison group met the same criteria as those in the intervention group. English language learners in the comparison group participated in the district's typical bilingual education program. In the primary grades, for approximately 20–30 minutes each day, English language learners received instruction in English language development. Instruction for the rest of the day was in Spanish. When they demonstrated proficiency in Spanish reading and writing, and basic proficiency in relation to oral English skills (the end of second grade through the beginning of third grade), they continued Spanish language arts, and began receiving reading and writing instruction in English (transition stage). After approximately 3–6 months, English language learners began a mainstream English program. The comparison group did not receive the transition program that incorporated 12 instructional components, including <i>Instructional Conversations</i> and <i>Literature Logs</i> , which were taught to English language learners in the intervention group.
Primary outcomes and measurement	English language learners' reading and English language development were assessed using the Comprehensive Test of Basic Skills (CTBS) and the Criteria for Addition of Reading in English (CARE). Reading outcomes were also assessed using performance assessments. Because most English language learners in grades 1–4 took the CTBS and performance assessment in Spanish, findings from these grades are not included in this report. Only findings for fifth-graders, who took the tests in English, are included in this report (see Appendix A2.1 and A2.2 for more detailed descriptions of outcome measures).
Teacher training	Saunders (1999) noted that the teachers involved in the study worked with project staff for the full five years during which it was implemented. No further information about teacher training was reported.

Appendix A2.1 Outcome measures in the reading achievement domain

Outcome measure	Description
Comprehensive Test of Basic Skills (CTBS), Reading Subtest	The CTBS is a standardized achievement test used to assess English language proficiency (as cited in Saunders, 1999). It includes subtests of reading and language. Performance on the CTBS was reported in terms of Norm Curve Equivalent (NCE) scores.
Performance assessments: Reading achievement	The author developed this assessment, which consists of six tasks: narrative comprehension (measures story comprehension), informational comprehension (measures how well English language learners understand informational text), writing communication (measures the lucidity of the information that English language learners provide to peers about story content), writing conventions (measures spelling, punctuation, capitalization, and usage), dictation (measures spelling, punctuation, and capitalization), and independent reading and library use, in which English language learners indicate the books and stories that they have chosen to read, and items borrowed from the library (as cited in Saunders, 1999). The assessments were administered in a small-group setting and scored by trained project staff and teachers under the direction of the author. The inter-rater reliability was 61% based on exact agreement and 97% based on adjacent agreement (agreement within one point with a five-point scale).
Factual comprehension	This measures required English language learners to recall factual details about a story. The measure was made up of 10 items (as cited in Saunders & Goldenberg, 1999).
Interpretive comprehension	This measure included five questions related to the interpretation of events that occurred in a story (as cited in Saunders & Goldenberg, 1999).

Appendix A2.2 Outcome measures in the English language development domain

Outcome measure	Description
Comprehensive Test of Basic Skills (CTBS), Language Subtest	The CTBS is a standardized achievement test used to assess English language proficiency (as cited in Saunders, 1999). It includes subtests of reading and language. Performance on the CTBS was reported in terms of Norm Curve Equivalent (NCE) scores.
Performance assessments: Writing achievement	The author developed this assessment, which consists of six tasks: narrative comprehension (measures story comprehension), informational comprehension (measures how well English language learners understand informational text), writing communication (measures the lucidity of the information that English language learners provide to peers about story content), writing conventions (measures spelling, punctuation, capitalization, and usage), dictation (measures spelling, punctuation, and capitalization), and independent reading and library use, in which English language learners indicate the books and stories that they have chosen to read, and items borrowed from the library (as cited in Saunders, 1999). The assessments were administered in a small-group setting and scored by trained project staff and teachers under the direction of the author. The inter-rater reliability was 61% based on exact agreement and 97% based on adjacent agreement (agreement within one point with a five-point scale).

Appendix A3.1 Summary of study findings included in the rating for the reading achievement domain¹

Outcome measure	Study sample	Sample size (students/schools)	Authors' findings from the study		Mean difference ³ (<i>IC</i> & <i>LL</i> – comparison)	WWC calculations		
			Mean outcome (standard deviation ²)			Effect size ⁴	Statistical significance ⁵ (at $\alpha = 0.05$)	Improvement index ⁶
			<i>IC</i> & <i>LL</i> group	Comparison group				
Saunders & Goldenberg, 1999 (randomized controlled trial) ⁷								
Factual comprehension	Grades 4–5	32/1	14.88 (3.03)	12.12 (3.50)	2.76	0.82	Statistically significant	+29
Interpretive comprehension	Grades 4–5	32/1	5.56 (2.73)	3.62 (1.96)	1.94	0.80	Statistically significant	+29
Average ⁸ for reading achievement (Saunders & Goldenberg, 1999)						0.81	Statistically significant	+29
Saunders, 1999 ⁹ (quasi-experimental design) ¹⁰								
Performance assessment	Grade 5	84/10	2.79 (0.65)	2.25 (0.47)	0.50	0.94	Statistically significant	+33
Comprehensive Test of Basic Skills (CTBS), Reading Subtest	Grade 5	84/10	41.90 (10.64)	34.51 (12.40)	7.39	0.63	ns	+24
Average ⁸ for reading achievement (Saunders, 1999)						0.79	Statistically significant	+28
Domain average ⁸ for reading achievement across all studies						0.80	Statistically significant	+29

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the average improvement index. Subgroup findings from the same studies are not included in these ratings, but are reported in Appendix A4.1.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are; a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect size calculation, see [Technical Details of WWC-Conducted Computations](#).
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and the percentile rank of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results.
7. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the [WWC Tutorial on Mismatch](#). See [Technical Details of WWC-Conducted Computations](#) for the formulas the WWC used to calculate statistical significance. In the case of Saunders & Goldenberg (1999), no correction for clustering was needed. The WWC did, however, correct for multiple comparisons.
8. The WWC-computed average effect sizes for each study and for the domain across studies are simple averages rounded to two decimal places. The average improvement indices are calculated from the average effect size.
9. Note that *Instructional Conversations* and *Literature Logs* were 2 of the 12 components of the language arts program that was implemented as part of the intervention. Therefore, findings were based on all 12 components, including *Instructional Conversations* and *Literature Logs*.
10. In the case of Saunders (1999), corrections for clustering and multiple comparisons were needed, which changed the statistical significance of outcomes on the CTBS.

Appendix A3.2 Summary of study findings included in the rating for the English language development domain¹

Outcome measure	Study sample	Sample size (students/schools)	Author's findings from the study		WWC calculations			
			Mean outcome (standard deviation ²)		Mean difference ³ (IC & LL – comparison)	Effect size ⁴	Statistical significance ⁵ (at $\alpha = 0.05$)	Improvement index ⁶
			IC & LL group	Comparison group				
Saunders, 1999 ⁷ (quasi-experimental design) ⁸								
Comprehensive Test of Basic Skills (CTBS), Language Subtest	Grade 5	84/10	45.38 (13.94)	37.80 (12.87)	7.58	0.56	ns	+21
Writing	Grade 5	84/10	2.87 (0.81)	2.37 (0.69)	0.50	0.66	ns	+24
Domain average ⁹ for English language development (Saunders, 1999)						0.61	ns	+23

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices. Subgroup findings from the same studies are not included in these ratings, but are reported in Appendix A4.2.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are; a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect size calculation, see [Technical Details of WWC-Conducted Computations](#).
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and the percentile rank of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results.
7. Note that *Instructional Conversations* and *Literature Logs* were 2 of the 12 components of the language arts program that was implemented as part of the intervention. Therefore, findings were based on all 12 components, including *Instructional Conversations* and *Literature Logs*.
8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the [WWC Tutorial on Mismatch](#). See [Technical Details of WWC-Conducted Computations](#) for the formulas the WWC used to calculate statistical significance. In the case of Saunders (1999), a correction for clustering was needed, which changed the statistical significance of outcomes on the CTBS.
9. This row provides the study average, which in this instance is also the domain average. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

Appendix A4.1 Summary of subgroup findings for the reading achievement domain¹

Outcome measure	Study sample	Sample size (students/schools)	Author's findings from the study		WWC calculations			
			Mean outcome (standard deviation ²)		Mean difference ³ (<i>IC & LL</i> – comparison)	Effect size ⁴	Statistical significance ⁵ (at $\alpha = 0.05$)	Improvement index ⁶
			<i>IC & LL</i> group	Comparison group				
Saunders, 1999 (quasi-experimental design) ⁷								
Comprehensive Test of Basic Skills (CTBS), Reading Subtest	Grade 5 (Transition 4) ⁸	56/10	40.56 (10.18)	33.62 (13.85)	6.94	0.56	ns	+21
Comprehensive Test of Basic Skills (CTBS), Reading Subtest	Grade 5 (Transition 5) ⁹	28/10	44.56 (11.41)	36.29 (9.04)	8.27	0.78	ns	+28

ns = not statistically significant

1. This appendix presents subgroup findings for measures that fall in the reading achievement domain. Total group scores were used for rating purposes and are presented in Appendix A3.1.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are; a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect size calculation, see [Technical Details of WWC-Conducted Computations](#).
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results.
7. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools (corrections for multiple comparisons were not done for findings not included in the overall intervention rating). For an explanation about the clustering correction, see the [WWC Tutorial on Mismatch](#). See [Technical Details of WWC-Conducted Computations](#) for the formulas the WWC used to calculate statistical significance. In the case of Saunders (1999), a correction for clustering was needed, which changed the statistical significance of outcomes on the CTBS.
8. These students began transitional English language arts (i.e., Transition II services) during the first semester of grade 4.
9. These students began transitional English language arts (i.e., Transition II services) during the first semester of grade 5.

Appendix A4.2 Summary of subgroup findings for the English language development domain¹

Outcome measure	Study sample	Sample size (students/ schools)	Authors' findings from the study		WWC calculations			
			Mean outcome (standard deviation ²)		Mean difference ³ (<i>IC</i> & <i>LL</i> – comparison)	Effect size ⁴	Statistical significance ⁵ (at $\alpha = 0.05$)	Improvement index ⁶
			<i>IC</i> & <i>LL</i> group	Comparison group				
Saunders & Goldenberg, 1999 (randomized controlled trial) ⁷								
Factual comprehension	Grades 4–5 (<i>Literature Logs</i> only versus comparison)	32/1	10.38 (3.72)	12.12 (3.50)	–1.74	0.47	ns	–18
Interpretive comprehension	Grades 4–5 (<i>Literature Logs</i> only versus comparison)	32/1	3.56 (1.86)	3.62 (1.96)	–0.06	0.03	ns	–1
Factual comprehension	Grades 4-5 (<i>Instructional Conversations</i> only versus comparison)	32/1	13.19 (4.10)	12.12 (3.50)	1.07	0.27	ns	+11
Interpretive comprehension	Grades 4-5 (<i>Instructional Conversations</i> only versus comparison)	32/1	4.88 (2.47)	3.62 (1.96)	1.26	0.55	ns	+21
Saunders, 1999 (quasi-experimental design) ⁸								
Comprehensive Test of Basic Skills (CTBS), Language Subtest	Grade 5 (Transition 4) ⁹	56/10	47.33 (15.63)	40.11 (12.81)	7.22	0.50	ns	+19
Comprehensive Test of Basic Skills (CTBS), Language Subtest	Grade 5 (Transition 5) ¹⁰	28/10	41.49 (8.99)	33.18 (12.13)	8.31	0.76	ns	+28
Writing	Grade 5 (Transition 4) ⁹	56/10	3.15 (0.78)	2.67 (0.60)	0.48	0.68	ns	+25
Writing	Grade 5 (Transition 5) ¹⁰	28/10	2.29 (0.54)	1.76 (0.40)	0.53	1.08	Statistically significant	+36

ns = not statistically significant

1. This appendix presents subgroup findings for measures that fall in the English language development domain. Total group scores were used for rating purposes and are presented in Appendix A3.2.

2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are; a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.

3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.

4. For an explanation of the effect size calculation, see [Technical Details of WWC-Conducted Computations](#).

Appendix A4.2 Summary of subgroup findings for the English language development domain *(continued)*

5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and the percentile rank of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results.
7. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools (corrections for multiple comparisons were not done for findings not included in the overall intervention rating). For an explanation about the clustering correction, see the [WWC Tutorial on Mismatch](#). See [Technical Details of WWC-Conducted Computations](#) for the formulas the WWC used to calculate statistical significance. In the case of Saunders & Goldenberg (1999), no correction for clustering was needed.
8. In the case of Saunders (1999), a correction for clustering was needed, but since the results were non-significant prior to the correction for clustering, there would have been no change in statistical significance after the correction was made.
9. These students began transitional English language arts (i.e., Transition II services) during the first semester of grade 4.
10. These students began transitional English language arts (i.e., Transition II services) during the first semester of grade 5.

Appendix A5.1 *Instructional Conversations and Literature Logs* rating for the reading achievement domain

The WWC rates the effects of an intervention in a given outcome domain as: positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.¹

For the outcome domain of reading achievement, the WWC rated *Instructional Conversations and Literature Logs* as having potentially positive effects. It did not meet the criteria for positive effects because there was no study that met the WWC evidence standards for a strong design. The remaining ratings (mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered, as the intervention was assigned the highest applicable rating.

Rating received

Potentially positive effects: Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

Met. In both studies on *Instructional Conversations and Literature Logs*, the intervention showed statistically significant positive effects.

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

Met. No studies showed a negative effect.

Other ratings considered

Positive effects: Strong evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a strong design.

Not met. The *Instructional Conversations and Literature Logs* intervention has two studies that examined reading achievement, both of which met WWC evidence standards with reservations and therefore did not qualify for a strong design.

- Criterion 2: No studies showing statistically significant or substantively important *negative* effects.

Met. No studies showed a negative effect.

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain level effect. The WWC also considers the size of the domain level effect for ratings of potentially positive or potentially negative effects. See the [WWC Intervention Rating Scheme](#) for a complete description.

Appendix A5.2 Instructional Conversations and Literature Logs rating for the English language development domain

The WWC rates the effects of an intervention in a given outcome domain as: positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.¹

For the outcome domain of English language development, the WWC rated *Instructional Conversations and Literature Logs* as having potentially positive effects. It did not meet the criteria for positive effects because there was no study that met the WWC evidence standards for a strong design. The remaining ratings (mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered, as the intervention was assigned the highest applicable rating.

Rating received

Potentially positive effects: Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

Met. The one study that examined the intervention's effect on English language development showed substantively important positive effects.

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

Met. No studies showed a negative effect.

Other ratings considered

Positive effects: Strong evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a strong design.

Not met. *Instructional Conversations and Literature Logs* had only one study that examined English Language Development, which met WWC evidence standards with reservations and thus did not qualify for a strong design.

- Criterion 2: No studies showing statistically significant or substantively important *negative* effects.

Met. No studies showed a negative effect.

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain level effect. The WWC also considers the size of the domain level effect for ratings of potentially positive or potentially negative effects. See the [WWC Intervention Rating Scheme](#) for a complete description.

(continued)